

New

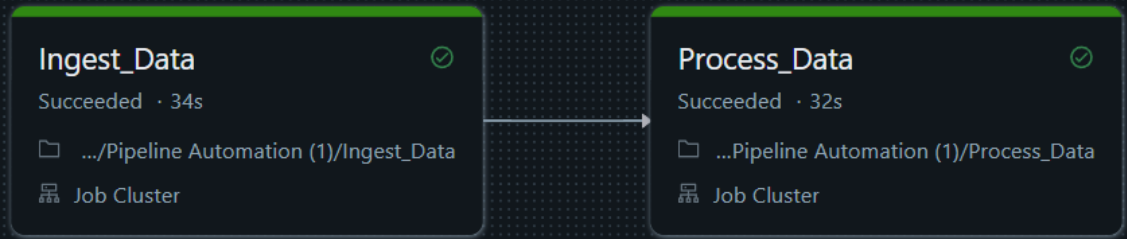
- Workspace
- Recents
- Catalog
- Workflows
- Compute
- Marketplace
- SQL
 - SQL Editor
 - Queries
 - Dashboards
 - Genie
 - Alerts
 - Query History
 - SQL Warehouses
- Data Engineering
 - Job Runs
 - Data Ingestion
 - Pipelines
- Machine Learning
 - Playground

Workflows > Jobs > Ingest_Process_Pipeline >

Ingest_Process_Pipeline run

Send feedback Delete job run Repair run

Graph Timeline List



Job run details

Job ID	697993054910077
Job run ID	316455260949517
Launched	Manually
Started	May 08, 2025, 08:48 AM
Ended	May 08, 2025, 08:49 AM
Duration	1m 8s
Queue duration	-
Status	Succeeded
Lineage	0 upstream tables, 3 downstream tables

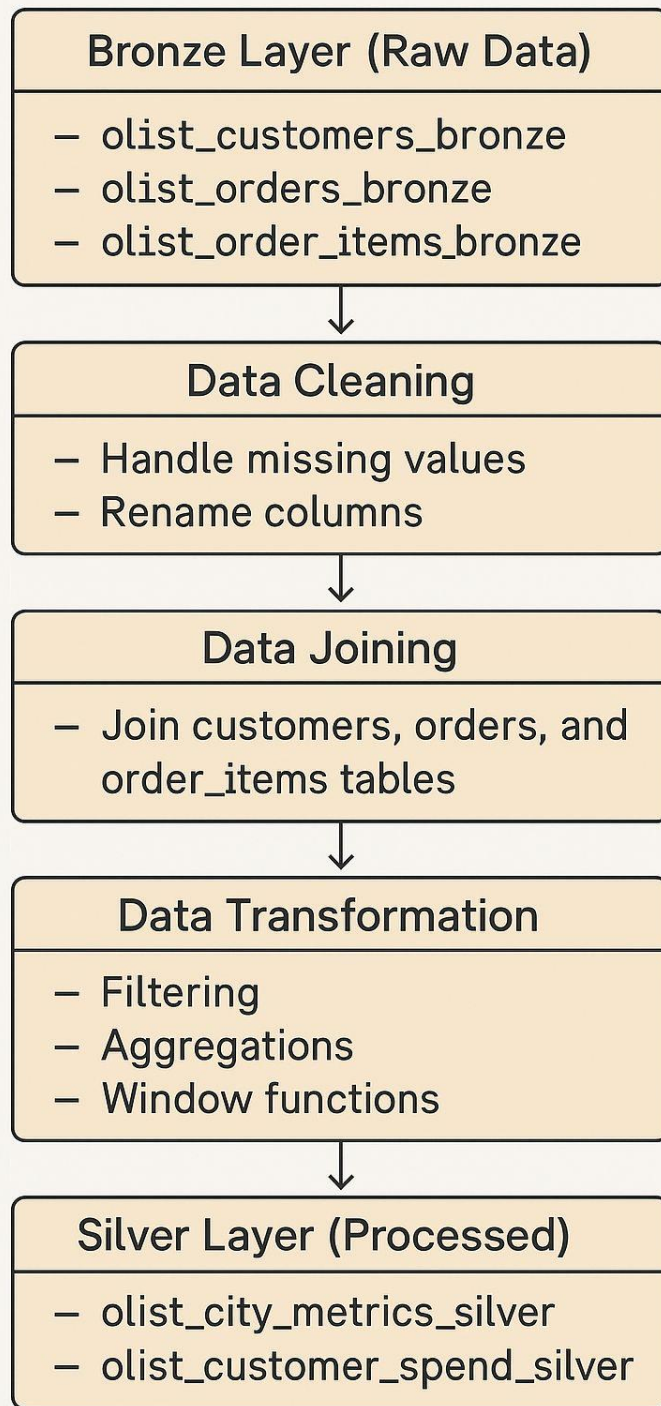
View run events

Compute

Job Cluster

Driver: Standard_D4s_v3 · Workers: Standard_D4s_v3 · 1-1 worker · Release: 13.3.36

View details Spark UI Logs Metrics



+ New

Workspace

Recents

Catalog

Workflows

Compute

Marketplace

SQL

SQL Editor

Queries

Dashboards

Genie

Alerts

Query History

SQL Warehouses

Data Engineering

Job Runs

Data Ingestion

Pipelines

Machine Learning

Workflows > Jobs >

Ingest_Process Pipeline ☆

Runs Tasks

ata

ne Automation (1)/Ing

er

Task name* ⓘ

Type*

Source* ⓘ

Path* ⓘ

Cluster* ⓘ

Schedules & Triggers

Trigger Status

☒ Active

☐ Paused

Trigger type

Scheduled

Schedule type

Simple

Advanced

Periodic ⓘ

Every

1

Day

Cancel

Save

ⓘ Jobs running on all-purpose clusters are considered all-purpose compute. [Learn more](#)

No job paramet

Edit paramet

Cancel

Save task

Process_Data

Python

Tabs: OFF



File

Edit

View

Run

Help

Last edit was 8 minutes ago

Run all



Workspace



← Pipeline Automation (1)



Ingest_Data



Process_Data



Silver_Table



3



4

```
from pyspark.sql.functions import sum, rank
from pyspark.sql.window import Window

spend_df = joined_df.groupBy("customer_id", "customer_city") \
    .agg(sum("freight_value").alias("total_spend"))

window_spec = Window.partitionBy("customer_city").orderBy(col("total_spend").desc())

ranked_df = spend_df.withColumn("rank", rank().over(window_spec))
```



5

```
agg_df.write.format("delta") \
    .mode("overwrite") \
    .saveAsTable("dev3.demodb.olist_city_metrics_silver")

ranked_df.write.format("delta") \
    .mode("overwrite") \
    .saveAsTable("dev3.demodb.olist_customer_spend_silver")
```